

ABSTRACT

A system and method for extracting a plurality of structured data from one or more information sources. The method comprises receiving the information sources, receiving at least one pattern descriptor selected from a graphical user interface, and 5 receiving one or more templates with each templates having at least one pattern descriptor. The method then proceeds to apply the one or more templates to the information sources. The method generates the plurality of structured data in a common format by parsing the information sources with the templates. The method stores the structured data in the common format.